

# Phosphate control of *phoA*, *phoC* and *phoD* gene expression in *Streptomyces coelicolor* reveals significant differences in binding of PhoP to their promoter regions

Alexander K. Apel,<sup>1,2</sup> Alberto Sola-Landa,<sup>1</sup> Antonio Rodríguez-García<sup>1</sup> and Juan F. Martín<sup>1,2</sup>

Correspondence  
Juan Francisco Martín  
jf.martin@unileon.es

<sup>1</sup>Instituto de Biotecnología de León, INBIOTEC, Parque Científico de León, Av. Real 1, 24006 León, Spain

<sup>2</sup>Área de Microbiología, Fac. CC. Biológicas y Ambientales, Universidad de León, Campus de Vegazana, s/n, 24071, León, Spain

Three putative alkaline phosphatase genes, *phoA*, *phoC* and *phoD*, were identified in the genome of *Streptomyces coelicolor* by homology with the amino acid sequence obtained from the PhoA protein of *Streptomyces griseus*. PhoA and PhoC correspond to broad-spectrum alkaline phosphatases whereas PhoD is similar to a Ca<sup>2+</sup>-dependent phospholipase D of *Streptomyces chromofuscus*. The *phoA* and *phoD* genes were efficiently expressed in R5 medium under phosphate-limited conditions, as shown by studies using the *xyIE* reporter gene, whereas *phoC* was poorly transcribed under the same conditions. Expression of *phoA* was clearly PhoP-dependent since it was not transcribed in the *S. coelicolor*  $\Delta$ *phoP* mutant and was strongly activated under low phosphate concentrations. Similarly, expression of *phoD* was PhoP-dependent and highly sensitive to phosphate availability. By contrast, expression of *phoC* was not PhoP-dependent. Electrophoretic mobility shift assays showed that PhoP binds to the *phoA* and *phoD* promoters, but not to that of *phoC*. Footprinting studies with GST–PhoP revealed the presence of a PHO box (two direct 11 nt repeats) in the *phoA* promoter and two PHO boxes in the promoter of *phoD*. The transcription start points of the three promoters were identified by primer extension. The transcription start point of *phoD* coincides with the G of its translation start codon, indicating that this gene is transcribed as a leaderless mRNA. The deduced –10 and –35 regions of *phoD* (but not those of *phoA*) overlapped with the PHO boxes in this promoter, suggesting that an excess of PhoP interferes with binding of the RNA polymerase to this promoter. In summary, the three promoters showed clear differences in the modulation of their expression by PhoP.

Received 15 February 2007  
Revised 21 June 2007  
Accepted 26 June 2007

## INTRODUCTION

*Streptomyces* species are soil dwelling micro-organisms that secrete many extracellular enzymes to degrade polymeric substrates and to utilize the nutrients obtained after their hydrolysis (García-González *et al.*, 1991; Fernández-Ábalos *et al.*, 1992; Schaerlaekens *et al.*, 2004). Phosphate is an essential nutrient for growth of all micro-organisms. *Streptomyces* species and other soil micro-organisms may use inorganic phosphate, which occurs frequently in the form of partially insoluble phosphate salts, and organic phosphate, which is present in a variety of plant materials such as phytic acid (Martínez-Domínguez *et al.*, 2002), sugar phosphates and nucleotides (Martín & Demain,

1977). To obtain phosphate from organic compounds, *Streptomyces* and other bacteria use an extracellular alkaline phosphatase encoded by the *phoA* gene (Moura *et al.*, 2001). When the *Streptomyces griseus phoA* gene was cloned and compared with the genome of *Streptomyces coelicolor*, two other putative phosphatase genes (hereafter named *phoC* and *phoD*; *phoB* was avoided because this designation is used for the response regulator in *Escherichia coli*), in addition to the *phoA* orthologue, were found (Moura *et al.*, 2001; Bentley *et al.*, 2002). Different phosphatases may work on different phosphorylated substrates.

Species of the genus *Streptomyces* are the major producers of secondary metabolites, including antibiotics and other pharmacologically active compounds, pigments, toxins, plant growth factors, etc. (von Döhren & Gräfe, 1997; Demain & Fang, 2000; Martín *et al.*, 2000). The

Abbreviations: EMSA, electrophoretic mobility shift assay; DBD, DNA-binding domain; GST, glutathione S-transferase.

biosynthesis of most of these secondary metabolites is negatively controlled by phosphate and their production occurs only under phosphate-limited conditions (Martín & Demain, 1980; Martín *et al.*, 1994).

Expression of phosphate-regulated genes is modulated by the two-component system PhoR–PhoP (Sola-Landa *et al.*, 2003; Ghorbel *et al.*, 2006). Binding of the response-regulator PhoP to the promoters of phosphate-regulated genes was shown to modulate expression of primary metabolism genes, including *phoA*, by Western blot analysis (Sola-Landa *et al.*, 2003), and genes involved in actinorhodin and undecylprodigiosin biosynthesis. Binding of the response regulator PhoP to the promoter regions of three genes of the *pho* regulon, *pstS*, *phoU* and *phoRP*, was shown both in *S. coelicolor* (Sola-Landa *et al.*, 2005) and *Streptomyces natalensis* (Mendes *et al.*, 2007), but direct binding of PhoP to the promoters of secondary metabolite biosynthesis genes has not been reported (Martín, 2004).

Synthesis of the PhoA alkaline phosphatase, as shown by Western blot analysis, is regulated positively by PhoP since in deletion mutants (either  $\Delta phoP$  or  $\Delta phoRP$ ) this protein is not detected (Sola-Landa *et al.*, 2003). However, it is not known whether the two-component system exerts its action by direct binding of PhoP to specific sequences in the *phoA* promoter region or whether regulation proceeds through other DNA-binding proteins by a cascade mechanism. Moreover, it is not known if the different phosphatase genes are regulated or not by phosphate availability or whether they respond to the same or to different type(s) of phosphate control.

It was, therefore, important to perform a comparative study of the mechanism of phosphate control of *phoA*,

*phoC* and *phoD* expression using DNA electrophoretic mobility shift assays (EMSAs) as well as reporter gene expression.

In this article we report significant differences in the control of the three putative phosphatase-encoding genes. EMSA and footprinting analyses provided evidence for the presence of PhoP-binding sequences (PHO boxes) in the promoters of *phoA* and *phoD* but not in that of *phoC*, which is regulated by a different mechanism.

## METHODS

**Bacterial strains and plasmids.** The bacterial strains and plasmids used in this work are listed in Table 1. Oligonucleotides were designed to amplify 500 bp fragments of the promoters of *phoA*, *phoC* and *phoD* containing sites for the restriction enzymes *Hind*III in the forward (F) and *Bgl*II or *Bam*HI in the reverse (R) primers (underlined in the sequences below).

The *phoA* promoter was cloned from cosmid SCC75A by PCR with the primer pair F1-TAATAAGCTTCCCCTTGAGGTTGGTGTC and R1-ATAGATCTGTTCCCTCCCCTTGCGG, resulting in a 517 bp fragment containing the promoter and the first two bases of the coding sequence. The *phoC* promoter was cloned from cosmid SCF43A by PCR with the primer pair F2-TAATAAGCTTCACCGCTGCCGTCGTA and R2-TGGGATCCATCTGGTGTCCCTCTCGTA, resulting in a 532 bp fragment containing the promoter and the first 10 bases of the coding sequence. The *phoD* promoter was cloned from cosmid SC4G6 by PCR with the primer pair F3-TAATAAGCTTCAAGACCGCGCTGTACTCCACGAA and R3-TAGGATCCAGCGGGGCGGCGAGCAC, resulting in a 529 bp fragment containing the promoter and the first 98 bases of the coding sequence (a region important for expression of this gene; see Results). The different nucleotide distance from the fusion site to the translation start codon of the *xylE* gene for the *phoD* promoter as compared to those of *phoA* and *phoC* might affect the intensity of expression from the different promoters.

**Table 1.** Bacterial strains and plasmids

Strain/plasmid	Characteristics	Reference
<b>Strains</b>		
<i>S. lividans</i> JI 1326	Wild-type	John Innes Centre*
<i>S. coelicolor</i> M145	Wild-type	Kieser <i>et al.</i> (2000)
<i>S. coelicolor</i> INB101	<i>S. coelicolor</i> M145 $\Delta phoP$	Rodríguez-García <i>et al.</i> (2007)
<i>E. coli</i> DH5 $\alpha$	F' $\phi 80$ <i>dlacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U16 <i>recA1 endA1 hsdR17</i> (r <sup>-</sup> m <sup>+</sup> ) <i>supE44 med</i> $\lambda^-$ <i>thi-1 gyrA relA1</i>	Hanahan (1983)
<b>Plasmids</b>		
pGEM-T Easy	Vector system for the cloning of PCR products	Promega
pGEM-P <i>phoA</i>	PCR product carrying the <i>phoA</i> promoter cloned into pGEM-T Easy	This work
pGEM-P <i>phoC</i>	PCR product carrying the <i>phoC</i> promoter cloned into pGEM-T Easy	This work
pGEM-P <i>phoD</i>	PCR product carrying the <i>phoD</i> promoter cloned into pGEM-T Easy	This work
pGEM-P <i>phoAp</i>	pGEM-P <i>phoA</i> minus the <i>Mlu</i> I– <i>Bam</i> HI fragment	This work
pGEM-P <i>phoDp</i>	pGEM-P <i>phoD</i> minus the <i>Mlu</i> I– <i>Mlu</i> I fragment	This work
pIJ4083	Promoter-probe plasmid using <i>xylE</i> as reporter	Clayton & Bibb (1990)
pIJ4083-P <i>phoA</i>	<i>phoA</i> promoter cloned into pIJ4083	This work
pIJ4083-P <i>phoC</i>	<i>phoC</i> promoter cloned into pIJ4083	This work
pIJ4083-P <i>phoD</i>	<i>phoD</i> promoter cloned into pIJ4083	This work

\*Collection of micro-organisms of the John Innes Centre, Colney Lane, Norwich NR4 7UH, UK.

After PCR amplification the promoters were cloned into pGEM-T Easy (Promega). All the promoters were sequenced using pUC/M13 forward and reverse primers in a Perkin Elmer ABI PRISM 310 Genetic Analyzer, extracted from the plasmid via digestion with *HindIII/BglI* or *HindIII/BamHI* and ligated into pIJ4083 in order to construct transcriptional fusions with the *xylE* reporter gene (Clayton & Bibb, 1990). The obtained promoter-probe plasmids pIJ4083-*PphoA*, pIJ4083-*PphoC* and pIJ4083-*PphoD* were first transformed into *Streptomyces lividans*, then extracted, confirmed by restriction with *HindIII* and *XbaI* and transformed into *S. coelicolor*.

**Promoter expression studies.** Promoter expression was studied using the *xylE* reporter gene of plasmid pIJ4083. Approximately  $3 \times 10^6$  spores were pregerminated in  $2 \times$  TY medium (Kieser *et al.*, 2000) for 8 h at 30 °C and 250 r.p.m. After centrifugation they were resuspended in 100 µl R5 medium (Kieser *et al.*, 2000) and used to inoculate 100 ml R5 medium supplemented with high (1.8 mM  $\text{KH}_2\text{PO}_4$ ) or low (0.009 mM  $\text{KH}_2\text{PO}_4$ ) concentrations of added phosphate. Thiostrepton was used as selective antibiotic at a final concentration of  $5 \mu\text{g ml}^{-1}$ . Samples were taken from two flasks at 18, 24, 27, 30, 33, 36, 39 and 48 h and catechol-2,3-dioxygenase activity was measured in triplicate as described by Kieser *et al.* (2000).

**DNA-protein binding assays.** To detect binding of PhoP to the promoters, studies were performed by EMSAs using glutathione S-transferase (GST)-PhoP and GST-PhoP<sup>DBD</sup> (DNA-binding domain) fusion proteins as described by Sola-Landa *et al.* (2005). The promoters were excised from the pGEM-T Easy plasmid with restriction enzymes and labelled at both ends with digoxigenin using the DIG Oligonucleotide 3'-Labelling kit, 2nd generation (Roche Applied Science).

Samples were loaded onto a 5% polyacrylamide native gel (29:1) in 0.5% TBE buffer and electrophoresis was run for 5 h at 80 V. Afterwards the gel was electroblotted onto a nylon membrane for 1 h at 200 mA in 0.5% TBE buffer and the DNA was fixed by UV cross-linking to the membrane. Labelled DNA was detected with anti-digoxigenin antibodies by chemoluminescence development with CDP-Star reagent (Roche Applied Science).

**Footprinting assays.** To obtain small promoter fragments, needed for footprinting studies, a fragment was deleted from the plasmids by digestion with *BamHI/MluI* in the case of pGEM-*PphoA* and with *MluI* alone in the case of pGEM-*PphoD*, using restriction sites located in the promoter and the flanking plasmid region but maintaining the annealing sequences for M13-20 and reverse primers. The plasmids obtained by religation were used to amplify the promoter fragments of *phoA* (*PphoAp*) and *phoD* (*PphoDp*) for DNase I footprinting studies using M13-20 and reverse primers, one of them fluorescein-labelled. The assays were performed by the fluorescein labelling procedure as described by Rodríguez-García *et al.* (1997) using the GST-PhoP<sup>DBD</sup> protein as described by Sola-Landa *et al.* (2005). After amplification the PCR products were purified by agarose-gel electrophoresis and the DNA concentrations were determined by spectrophotometry with a GeneQuant spectrophotometer (Amersham Biosciences).

The reaction components were the same as in the DNA binding studies. Labelled DNA fragments (0.28 pmol) and GST-PhoP<sup>DBD</sup> protein were added to a final volume of 28 µl, and incubated at 30 °C for 30 min. Lyophilized bovine pancreas DNase I (Roche grade I) was reconstituted in 20 mM Tris/HCl, pH 7.0, 50 mM NaCl,  $100 \mu\text{g ml}^{-1}$  BSA, 1 mM DTT and 50% glycerol, to a final concentration of 20 units  $\mu\text{l}^{-1}$ . Further dilutions were made in the same solution supplemented with 10% glycerol. Nuclease digestions were carried out with 2 µl of the 1:8000 dilution ( $5 \times 10^{-3}$  units) at 25 °C for 1 min and stopped with 120 µl 40 mM EDTA in 9 mM Tris/HCl, pH 8.0. After phenol/chloroform purification and ethanol precipitation, samples were loaded in an ALF DNA sequencer

(Amersham Biosciences) and compared with the corresponding sequencing reaction using the same labelled primer as used for the amplification. Results were analysed with the Fragment Manager program (Amersham Biosciences).

**RNA extraction and primer extension.** The *S. coelicolor* M145 derivatives containing pIJ4083-*PphoA*, pIJ4083-*PphoC* and pIJ4083-*PphoD* were grown under the same conditions as described for the expression studies up to the time of maximum expression of each promoter (36 h for *phoA* and *phoC*; 39 h for *phoB*; see Results). RNA was then isolated by the hot phenol procedure adapted to *Streptomyces* (Patek *et al.*, 2003). The transcriptional start sites were determined by primer extension using the 5'-fluorescein-labelled oligonucleotide O6-GCGATCGCTGCCACTGC as a primer (complementary to the 5' coding region of *xylE* from pIJ4083 vector). The reaction product was loaded in an ALF DNA sequencer (as described above) and analysed with the Fragment Manager program.

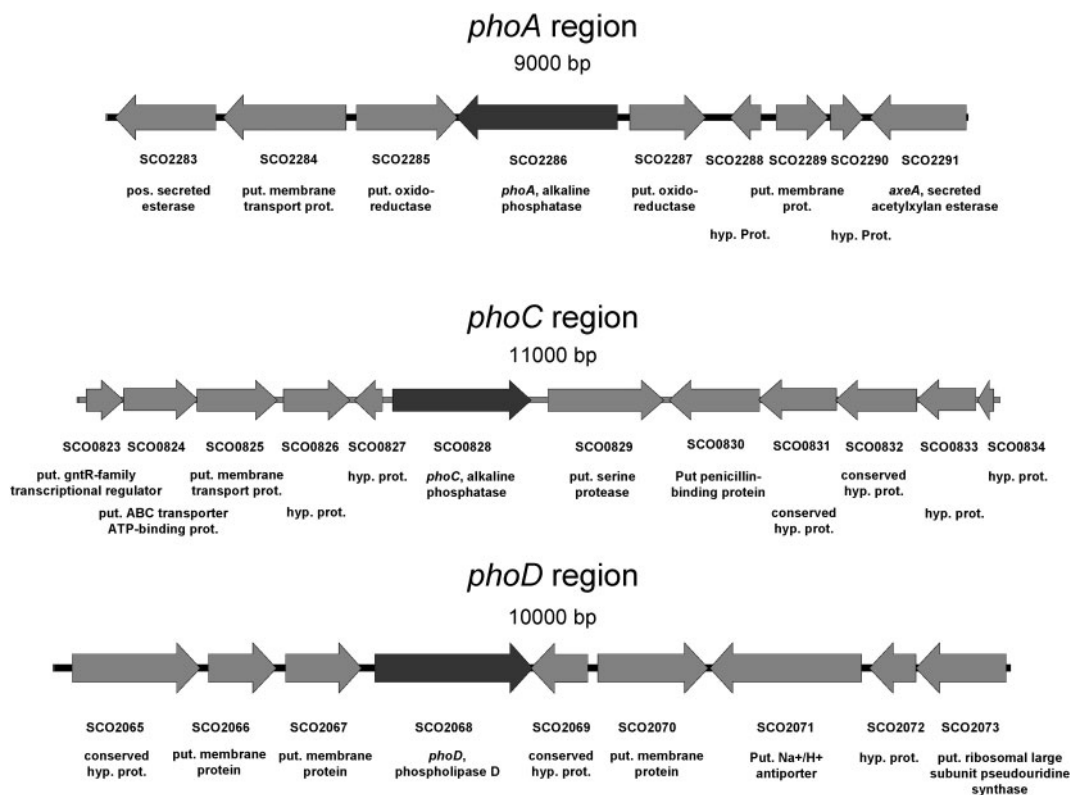
## RESULTS

### Three different *S. coelicolor* alkaline phosphatase genes: bioinformatic analysis indicates that PhoA, PhoC and PhoD are extracellular enzymes

Three putative alkaline phosphatase genes, *phoA*, *phoC* and *phoD*, were identified in the genome of *S. coelicolor* by a comparative search with the *phoA* gene of *S. griseus* (corresponding to the sequenced alkaline phosphatase protein; Moura *et al.*, 2001). The encoded proteins PhoA (558 aa, SCO2286), PhoC (553 aa, SCO0828) and PhoD (551 aa, SCO2068) (Fig. 1) showed amino acid identities of 73, 67 and 36%, respectively with *S. griseus* PhoA. Bioinformatic analysis of the three proteins using the PFAM and SYSTEMS programs revealed that all three contain an alkaline phosphatase domain. The *phoA*- and *phoC*-encoded proteins showed high similarity to the alkaline phosphatases of *Bacillus* species, whereas *phoD* encodes a protein with high similarity (85% identical residues) to a well-known  $\text{Ca}^{2+}$ -dependent phospholipase D precursor of *Streptomyces chromofuscus* (Stieglitz *et al.*, 2001; Yang & Roberts, 2002). This phospholipase catalyses two reactions: (i) hydrolysis of different phospholipids to generate phosphatidic acid, and (ii) transphosphatidylation with primary or secondary alcohols to generate different phospholipids. Therefore, *S. coelicolor* PhoD seems to correspond to a phospholipase D (see Discussion).

Since the total phosphatase activity in *S. coelicolor* is a mixture of the different phosphatases, expression of each gene was quantified separately by coupling it to the *xylE* reporter gene.

Alkaline phosphatases and related proteins may be extracellular enzymes involved in phosphate scavenging or they may play roles in secondary metabolite biosynthesis (e.g. streptomycin-6-phosphate phosphatase encoded by *strK* (Mansouri & Piepersberg, 1991) or in catabolic reactions (e.g. degradation of xenobiotics). Hydrophobicity plots using two different programs, TMpred and TMHMM, predicted all three proteins PhoA, PhoC and PhoD to be soluble proteins.



**Fig. 1.** Physical map of the *S. coelicolor* DNA regions around *phoA*, *phoC* and *phoD* (black arrows). The arrows indicate ORFs and the orientation of transcription.

PhoA (SCO2286) and PhoD (SCO2068) were shown to be secreted via the twin arginine translocation (Tat) pathway, a major route for protein secretion in *Streptomyces*, (Widdick *et al.*, 2006). Their N-terminal regions contain the conserved twin-arginine motif defined as R-R-X-Φ-Φ, where Φ represents a hydrophobic amino acid (Berks, 1996). Using the TatP algorithm (Bendtsen *et al.*, 2005) for prediction of proteins exported via the Tat pathway we identified a clear Tat motif in PhoC, with the sequence GAAARHLGRRRFLTVTAA (amino acids 23–40), nearly identical to the Tat motif RAAARSLGRRRFLTVTGA (amino acids 27–44) predicted for PhoA (Widdick *et al.*, 2006), suggesting that PhoC might be secreted via the Tat pathway.

Using neural networks and hidden Markov models trained on Gram-positive bacteria for protein secretion via the Sec or the SRP pathway, PhoA and PhoD are also predicted to be secreted. The cleavage site of the PhoA signal peptide is predicted to lie between amino acids 62 and 63 with a probability of 95.3%. The corresponding sequence is ASA/A (/ indicates the cleavage site). The cleavage site of PhoD is located (with a probability of 81.3%) between amino acids 44 and 45 in the sequence ADA/A. The overall probability for both proteins to be secreted is 95.5% for PhoA and 100% for PhoD. The above-mentioned models, which are trained to predict proteins secreted via the Sec or

the SRP pathway, indicate that PhoC is not secreted through these secretion pathways (with a secretion probability of only 30.5%).

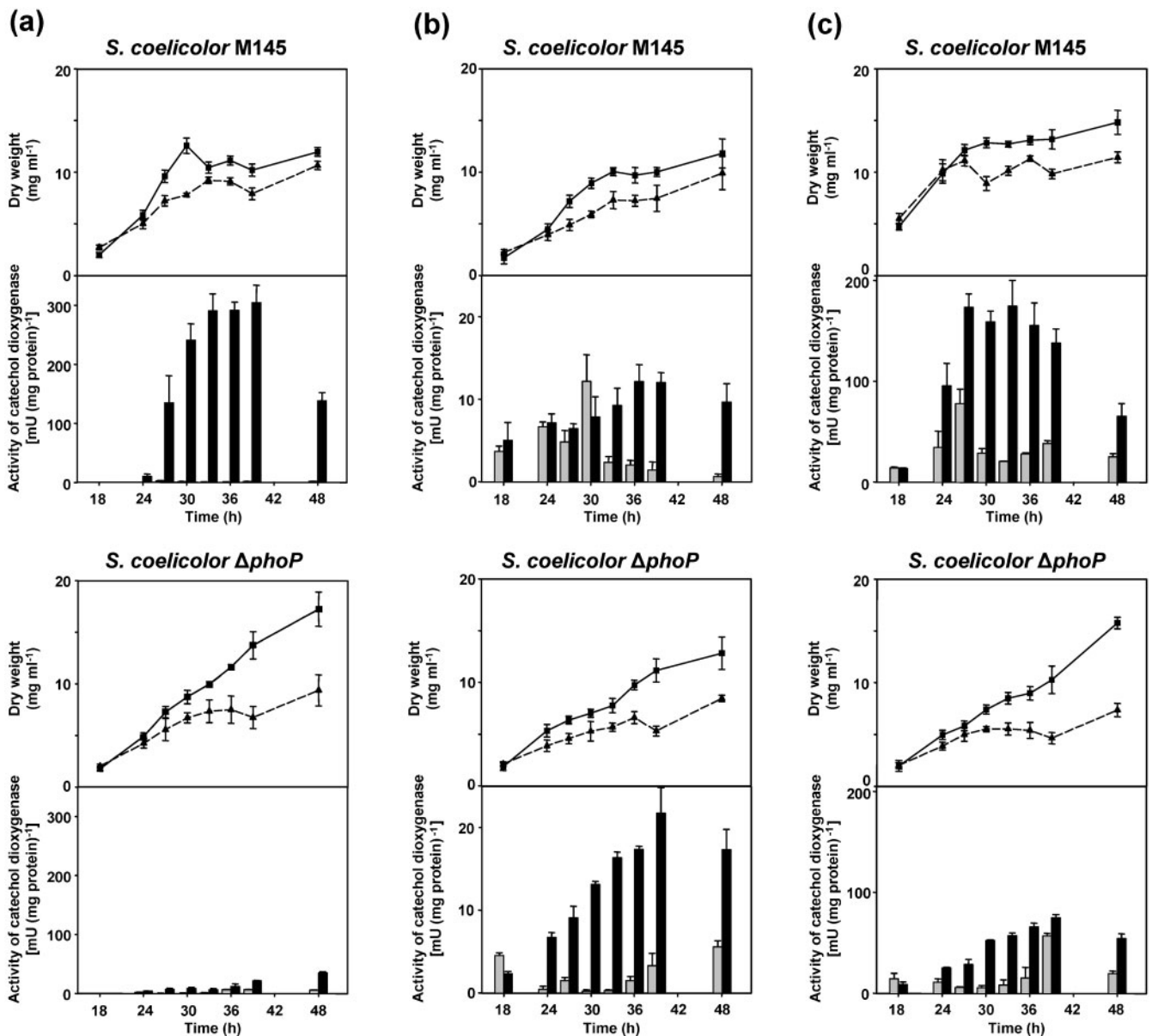
### Expression of *phoA* is strongly induced by phosphate limitation and this induction is PhoP-dependent

The promoter of *phoA* (*PphoA*) in *S. coelicolor* was weakly expressed in high phosphate but highly expressed in low-phosphate conditions (Fig. 2a). In phosphate-limited batch cultures *phoA* is efficiently expressed, as measured by quantification of reporter activity, and reached its maximum activity at about 39 h (Fig. 2a), decreasing thereafter.

Expression of *phoA* was nearly non-existent in the *S. coelicolor*  $\Delta$ *phoP* mutant, indicating that PhoP is strictly required for expression of this promoter, i.e. *phoA* expression is dependent upon activation by PhoP.

### The promoter of *phoC* shows a weak transcriptional activity and is not PhoP-dependent

Expression of *phoC* in wild-type *S. coelicolor* M145 is much lower than that of *phoA* measured using the same reporter system. The profile of expression of *phoC* was clearly



**Fig. 2.** Phosphate control of expression of *phoA* (a), *phoC* (b) and *phoD* (c) in the parental strain *S. coelicolor* M145 and the  $\Delta$ *phoP* mutant (upper and lower panels, respectively), using the promoters of each of the three genes coupled to the *xyIE* reporter gene. Cultures of each strain were performed in R5 medium supplemented with 0.009 mM phosphate ( $\blacktriangle$  and black bars) or with 1.8 mM phosphate ( $\blacksquare$  and grey bars). The growth curves of each transformant are shown in the upper part of each panel. The error bars indicate SD ( $n=6$ ). Note the different scales for XylE (catechol dioxygenase) activity for the three promoters.

different from that of *phoA*. It was induced under low-phosphate concentrations, activity reaching its maximum after 36–39 h in R5 phosphate-limited medium, and not decreasing significantly until after 48 h of growth (Fig. 2b). However, its expression was not maintained after 30 h in phosphate-supplemented medium, indicating that this gene might be subject to a complex dual regulation by phosphate. This gene does not encode a typical alkaline phosphatase (see Discussion) and its expression may be modulated differently in low- and high-phosphate conditions.

The expression of *phoC* is negatively regulated by phosphate in both the wild-type and the  $\Delta$ *phoP* mutant. In conditions of phosphate sufficiency the expression of *phoC* is only active in early growth (up to 30 h) then abruptly decreases. This early expression of *phoC* is not observed in the *phoP* mutant. A similar observation was made in the case of *phoD*. In order to explain these observations, we can argue that *phoC* and *phoD* are expressed during the exponential growth phase in high-phosphate conditions by a transcription factor that

depends on the presence of PhoP. The later expression of *phoC* observed under phosphate-limited conditions is not under the positive control of PhoP, since its expression is even higher in the *phoP* mutant strain; this indicates the existence of an unknown phosphate-responsive regulator, probably detecting intracellular phosphate limitation, which is more severe in the case of the *phoP* mutant due to its inability to activate the expression of genes of the PHO regulon.

### Expression of *phoD* is highly regulated by phosphate and partially PhoP-dependent

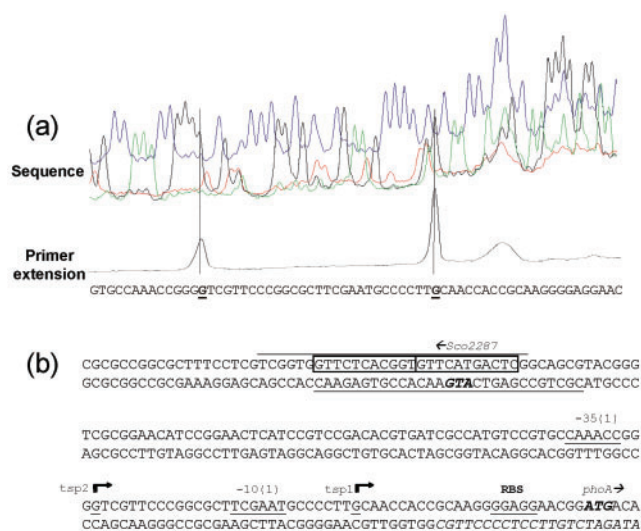
The expression of *phoD* increased under low-phosphate conditions, reaching an expression level nearly as high as that of *phoA*. Its expression was also significantly reduced by 1.8 mM phosphate.

In the  $\Delta$ *phoP* mutant the overall *phoD* expression decreased significantly (but was not completely suppressed) when compared to wild-type strain M145, indicating that although PhoP is clearly involved in the phosphate control of *phoD*, its expression is not as strictly PhoP-dependent, as observed for *phoA* (Fig. 2c). In conditions of phosphate limitation, the expression of *phoD* is very active throughout growth and this expression is clearly dependent upon PhoP. In conditions of phosphate sufficiency, the expression of *phoD* is active in early growth then abruptly decreases as for *phoC*; this early expression is not detected in the  $\Delta$ *phoP* mutant whereas a late expression (40 h) is observed. Our results suggest that another regulatory element besides PhoP plays a role in *phoD* induction under conditions of phosphate limitation.

### Characterization of the promoters by primer extension

The transcription start points of the *phoA*, *phoC* and *phoD* genes were determined by primer extension. In all three cases the origins of transcription were identified and, once the +1 sites were known, the corresponding -10 and -35 boxes of each promoter were established by comparison to the matrices reported by Bourn & Babb (1995) for *Streptomyces* that take into account the nucleotides occurring in 13-nucleotide stretches, including the -10 or -35 consensus hexamers, as described by Sola-Landa *et al.* (2005).

Two different transcription start points (*tsp1* and *tsp2*) were found upstream of the *phoA* open reading frame (Fig. 3). The first and stronger one was located at a guanosine 25 bp upstream of the start codon and corresponds to a -10 box (TCGAAT) and -35 box (CAAACC) separated by 18 bp. For the second transcription start point, located 53 bp upstream of the start codon, no consensus -10 and -35 boxes could be identified, suggesting that this second transcription start point is probably recognized by a different sigma factor. A putative RBS with the consensus core sequence GGAGG was found



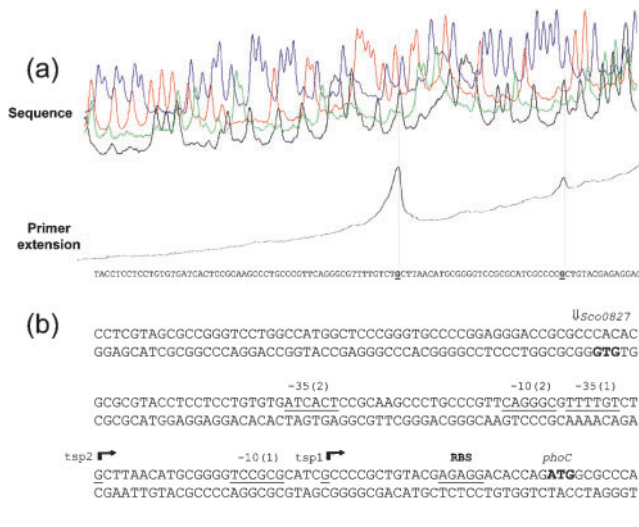
**Fig. 3.** Primer extension analysis of the *phoA* promoter (a) and nucleotide sequence of its promoter region (b). Two peaks were found in the primer extension analysis using 50  $\mu$ g RNA. The start codon of *phoA* and of its upstream gene are indicated in bold. The transcription start points *tsp1* and *tsp2* are indicated by bent arrows; the -10 and -35 region and the putative RBS are underlined. The region protected by PhoP in DNase footprinting assays is over- and underlined in the corresponding strand. The PHO sequence is indicated by boxes.

6 bp upstream of the ATG translation start initiation codon of *phoA*.

Transcription of *phoC* initiates from two different transcription start points (*tsp1* and *tsp2*) located 25 bp and 50 bp upstream of the start codon (Fig. 4). For both sites possible -10 and -35 boxes with the sequences -10 TCCGCG/-35 TTTTGT (*tsp1*) and -10 CAGGGC/-35 ATCACT (*tsp2*) separated by 17 bp (*tsp1*) and 18 bp (*tsp2*) could be identified. A possible RBS with the core sequence AGAGG is located 7 bp upstream of the start codon of *phoC*.

Transcription of *phoD* initiates at the first guanosine of its GTG translation start codon, for which no RBS was identified. A second GTG codon is found 69 bp downstream, but comparison of PhoD with the homologous proteins from other *Streptomyces* sequences shows high conservation within the first 23 amino acids and as those include the Tat secretion signal predicted by Widdick *et al.* (2006), translation is likely to start at the first GTG (Fig. 5). Its -10 box (TAACCT) and -35 box (CCCACT) are separated by 18 bp and overlap with the second protected region and the PHO box repeats located in it (Fig. 5).

All transcription initiation points were consistent with the identified -10 and -35 regions. We cannot exclude, however, other minor transcription start points that do not show up when using the primer extension technique.



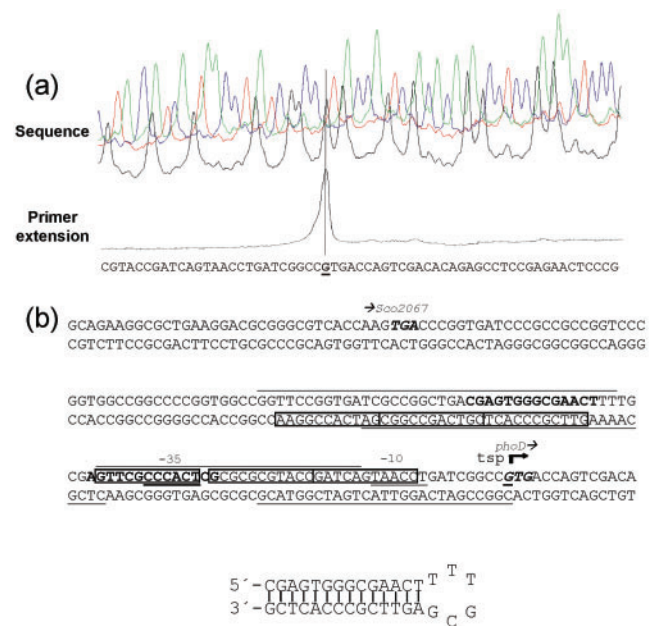
**Fig. 4.** Primer extension analysis of the *phoC* promoter (a) and nucleotide sequence of its promoter region (b). Two peaks were found in the primer extension analysis, which was carried out with 200  $\mu$ g RNA. The start codon of *phoC* and of its upstream gene are indicated in bold. The transcription start points are shown by bent arrows; the  $-10$  and  $-35$  region and the putative RBS are underlined. No consensus PHO boxes are evident in this promoter.

Upstream of the *phoD* gene there are three ORFs transcribed in the same orientation (Fig. 1); one of them encodes a conserved hypothetical protein (SCO2065) and the other two encode putative membrane proteins (SCO2066/SCO2067). Downstream of *phoD* there is a conserved hypothetical protein arranged in the opposite orientation. In the expression studies we observed promoter activity for the DNA region upstream of *phoD* (Fig. 2). The possibility that *phoD* is additionally transcribed in an operon together with these three genes seems unlikely because of the 14 bp palindromic sequence located between the transcription start point of *phoD* and SCO2067 (Fig. 5) that could form a stable stem-loop structure in the RNA and function as a transcriptional terminator.

### PhoP binds to the promoters of *phoA* and *phoD*, but not to that of *phoC*

The interaction of PhoP with the three promoters was tested by EMSAs using the *S. coelicolor* GST-PhoP fusion protein as well as a truncated version of PhoP containing only its DNA-binding domain (C-terminal region; GST-PhoP<sup>DBD</sup>). The PhoP<sup>DBD</sup> fragment shows constitutive binding to PHO boxes, independent of its phosphorylation status (Sola-Landa *et al.*, 2005).

The promoters of *phoA* (Fig. 6a) and *phoD* (Fig. 6c) showed a clear shift of the labelled DNA band. In both cases the binding was stronger with the truncated PhoP<sup>DBD</sup> than with the complete PhoP protein. In the EMSAs performed with the *phoD* promoter up to four bands were



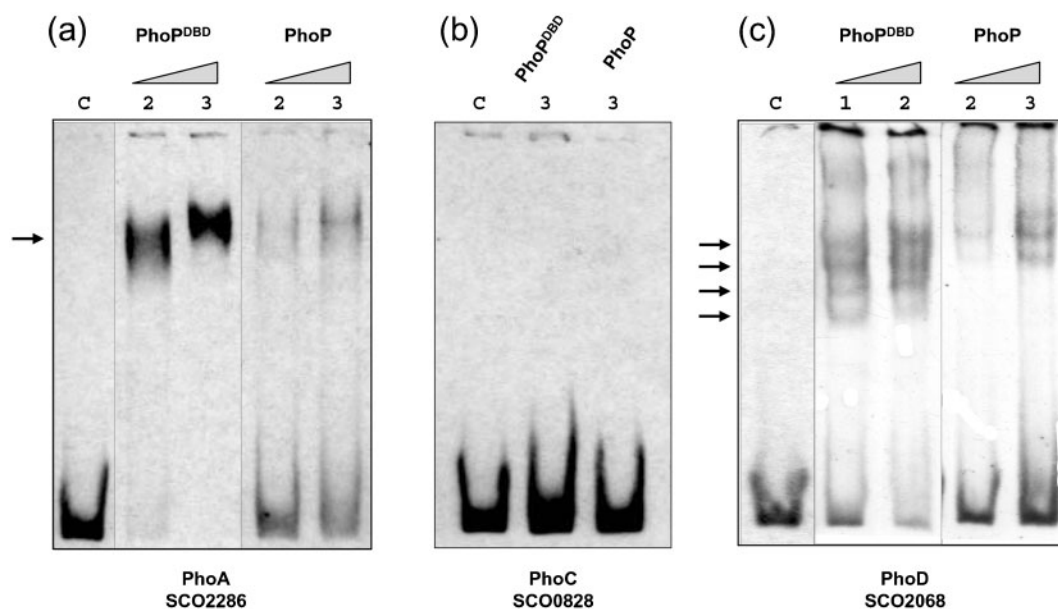
**Fig. 5.** Primer extension analysis of the *phoD* promoter (a) and nucleotide sequence of its promoter region (b). One clear transcription start point could be identified by primer extension carried out with 50  $\mu$ g RNA. The start codon of *phoD* and the stop codon of its upstream gene are indicated in bold. The transcription start point is indicated by a bent arrow; the  $-10$  and  $-35$  regions are underlined. The region protected by PhoP in DNase footprinting assays is over- and underlined in the corresponding strand; PHO-sequences are indicated by boxes. A 14 bp palindromic sequence, probably working as a transcriptional terminator for the upstream gene, is indicated in bold. The stem-loop structure of the putative terminator is shown below the promoter sequence.

observed, indicating that several DNA-protein complexes were formed due to binding of increasing amounts of PhoP<sup>DBD</sup> to the three Pho boxes existing in this promoter (see below), whereas the promoter of *phoA* formed only one shifted band.

An interesting finding was that the promoter of *phoC* did not show a shift reaction with either the complete PhoP or the truncated PhoP<sup>DBD</sup> version (Fig. 6b) under the same conditions as used for EMSA that showed mobility shifts for the *phoA* and *phoD* promoters. These results agree with the expression studies (Fig. 2b) and suggest that the *phoC* promoter lacks adequate PHO boxes. Computer analysis of the *phoC* promoter confirmed the absence of PHO boxes in this promoter, supporting the observed lack of binding of the PhoP response regulator to this DNA region.

### Footprinting assays of *phoA* and *phoD*: identification of PHO boxes in these promoters

To locate the PHO box repeat units in the promoters of *phoA* and *phoD*, footprinting assays were performed using



**Fig. 6.** EMSAs of the promoter regions of *phoA* (a), *phoC* (b) and *phoD* (c) using different concentrations of GST-PhoP<sup>DBD</sup> and GST-PhoP. Lanes: C, Control without protein; 1, 6.25 pmol; 2, 12.5 pmol; 3, 25 pmol protein. Arrows indicate shifted bands.

the PhoP<sup>DBD</sup> protein to protect them against DNase I digestion.

Results of the footprinting studies identified in the *phoA* promoter a PhoP-protected region of 29 bp in the coding strand and of 29 bp in the complementary strand, overlapping by 23 bp. In this region one complete Pho box containing two repeats of 11 bp (Sola-Landa *et al.*, 2005) with the sequence GTTCTCACGGT GTTCATGACTC was identified (the first nucleotide of each repeat is underlined) (Fig. 7). This sequence is located in the coding strand 103 or 75 bp upstream of the experimentally determined transcription start points *tsp1* and *tsp2*, respectively (see below).

In the *phoD* promoter two protected regions of 38 bp and 28 bp (separated by 6 bp) in the coding strand and 33 bp and 27 bp (separated by 16 bp) in the complementary strand were identified, overlapping by 27 bp and 11 bp, respectively (Fig. 8).

The upstream protected region contains three direct 11 nt repeats located in the complementary strand with the sequence GTTCGCCCACT CGTCAGCCGGC GATCACCGGAA (the first nucleotide of each repeat is underlined). The downstream protected region contains three direct repeats with the sequence GTTCGCCCACT C GCGCGCGTACC GATCAGTAACC located in the coding strand and separated by one nucleotide (C) between the first and the following two repeats (Fig. 8) (see Discussion).

The Pho box repeats in *phoD* are located 10 bp from the experimentally determined transcription start point,

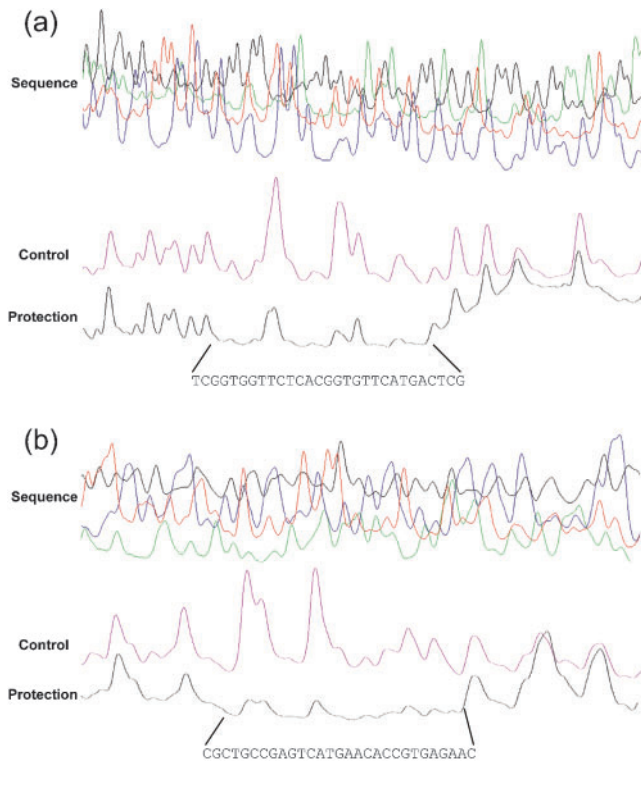
overlapping with the predicted -10 and -35 boxes (see below).

## DISCUSSION

Three different alkaline phosphatase genes were initially described in the *S. coelicolor* genome (Moura *et al.*, 2001; Bentley *et al.*, 2002). Interestingly, similar BLAST searches in the genomes of *Streptomyces avermitilis* and *Streptomyces scabies* only identified two possible alkaline phosphatases, corresponding to *phoA* and *phoD*; so the *phoC* gene of *S. coelicolor* seems to be a more recent addition to the *S. coelicolor* genome and might be involved in the biosynthesis of some *S. coelicolor*-specific metabolite. PhoA of *S. coelicolor* corresponds to an authentic alkaline phosphatase with high similarity to the well-known extracellular alkaline phosphatase of *S. griseus* (Moura *et al.*, 2001) and other micro-organisms. It has been shown to be secreted via the Tat pathway (Widdick *et al.*, 2006) but also contains a typical leader peptide as described for Sec-dependent secretion (Schaerlaekens *et al.*, 2004).

Expression of *phoA* is strongly activated by phosphate limitation (Fig. 2a) and is strictly dependent upon the binding of PhoP to the consensus PHO box identified in its promoter (see below). Transcriptional analysis revealed the existence of two transcription start sites, *tsp1* and *tsp2*, allowing the definition of two promoters. DNase I footprinting experiments revealed the presence of putative PHO boxes (two direct repeats of 11 bp) located 103 and 75 bp upstream of *tsp1* and *tsp2* respectively. These PHO boxes are in the





**Fig. 7.** Footprinting of the coding strand (a) and the complementary strand (b) of the *phoA* promoter region using GST-PhoP<sup>DBD</sup>. Shown is the sequencing reaction (above), a control without protein (middle) and the footprinting (below) using 1.06  $\mu$ M GST-PhoP<sup>DBD</sup>. The sequence of the protected region is given.

position of an activator site for both promoters that are thus both likely to be regulated by phosphate availability.

*phoC* encodes an alkaline phosphatase similar to PhoA, but its regulation pattern is clearly different from that of *phoA* and *phoD*. Its expression can be divided into two phases: an early PhoP-dependent phase up to 30 h and a later phase not directly dependent upon PhoP. Under conditions of phosphate limitation, the expression of *phoC* is active in early growth stages but keeps on increasing during later growth and is even higher in the *phoP* mutant.

Although the TatP program (Bendtsen *et al.*, 2005) predicts PhoC to be secreted via the Tat pathway, the KEGG database, a knowledge base for systematic analysis of gene functions (Kanehisa *et al.*, 2006), classifies PhoC as a phosphoric monoester hydrolase that may be involved in two pathways: (i)  $\gamma$ -hexachlorocyclohexane degradation and (ii) folate biosynthesis. Although both pathways occur in *S. coelicolor* (PATHS SCO00361 and SCO00790, respectively) there is no unequivocal evidence to conclude that PhoC corresponds to the phosphatase catalysing the dephosphorylation steps of these pathways.

The lack of binding of the PhoP response regulator to the *phoC* promoter suggests that the observed phosphate

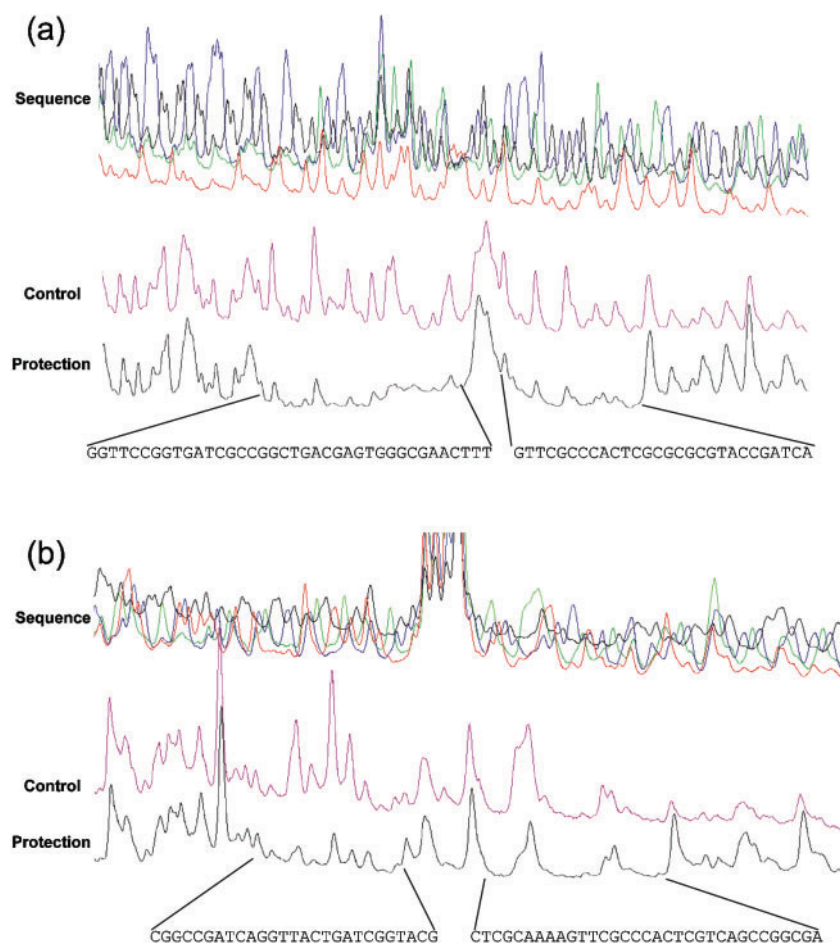
regulation of the *phoC* gene is not mediated by direct interaction of PhoP, in agreement with the lack of PHO boxes in this promoter region. Phosphate regulation of *phoC* during the early growth phase might involve an intermediary regulator whose expression might be under the control of PhoR/PhoP. The phosphate-response regulator PhoP binds to a significant number of phosphate-regulated genes in *Streptomyces* (Sola-Landa *et al.*, 2005), but other phosphate-controlled genes lack Pho boxes and are not recognized by PhoP (Rodríguez-García *et al.*, 2007).

The *S. coelicolor phoD* gene corresponds to a novel-type Ca<sup>2+</sup>-dependent phospholipase similar to that of *S. chromofuscus* (85% identical amino acid residues). These phospholipases of *S. coelicolor* and *S. chromofuscus* show little homology with other members of the phospholipase D superfamily. This PhoD phospholipase is also different from classical phospholipases described in other *Streptomyces* species. The PhoD proteins of *S. coelicolor* and *S. chromofuscus* lack the identifiable HKD motif common to members of the phospholipase D family. The *S. chromofuscus* enzyme has been characterized biochemically and shows phosphodiesterase activity; it is secreted and cleaved by an extracellular protease, increasing its enzymic activity (Geng *et al.*, 1999; Yang & Roberts, 2002). This enzyme binds to membrane phosphatidylcholine and facilitates vesicle aggregation and fusion (Stieglitz *et al.*, 1999, 2001).

Transcriptional analysis revealed the existence of a unique transcriptional start point coinciding with the first base of the translation initiation codon, allowing the definition of an unique promoter. The expression of *phoD* is activated under phosphate-limited conditions by PhoP, which binds to the two regions of 11 bp repeat units (PHO box) found in its promoter, overlapping partially with its predicted -10 and -35 boxes. The formation of up to four DNA-PhoP bands in the EMSA experiments correlates with the distinct organization of the PHO boxes in the promoter of this gene. The activation of *phoD* in the mutant strain *S. coelicolor*  $\Delta$ *phoP* (INB101) is weaker than that in the wild-type strain. This activation in the  $\Delta$ *phoP* mutant under phosphate-limitation conditions suggests that another regulatory factor besides PhoP plays a role in phosphate control.

The sequence of PHO boxes protected by binding of the PhoP protein was established recently in the promoters of the *pstS*, *phoRP* and *phoU* genes (Sola-Landa *et al.*, 2005; Rodríguez-García *et al.*, 2007). However, the number of validated PHO boxes in *Streptomyces* genomes is still very limited. The presence of two 11 bp direct repeats in the *phoA* promoter and six repeats in the *phoD* promoter that match the protected region allowed us to establish a new, better-fitting sequence for the Pho box as G(82%)G/T(85%)T(77%)C(90%)A(59%)Y(82%)Y(74%)Y(72%)R(62%)Y(64%)Y(74%).

Duplication of the number of PHO boxes in certain promoters is related to a potential modulation of the strong phosphate control (Mendes *et al.*, 2007). The



**Fig. 8.** Footprinting of the coding strand (a) and the complementary strand (b) of the *phoD* promoter region using GST-PhoP<sup>DBD</sup>. Shown is the sequencing reaction (above), a control without protein (middle) and the footprinting (below) using 1.06  $\mu\text{M}$  GST-PhoP<sup>DBD</sup>. The sequence of the protected region is given.

formation of up to four DNA-PhoP protein complexes in the binding reaction of the *phoD* promoter suggests that this promoter is subject to different degrees of PhoP control depending on the level of the phosphorylated response regulator in the cell.

## ACKNOWLEDGEMENTS

This work was supported by grants of the CICYT (BIO2003-01489), Madrid (Spain) and the European Union (Project ACTINOGEN, LSHM-CT-2004-005224) Brussels. A. K. Apel received a F.P.U. fellowship of the Ministry of Education and Science. We thank J. Merino, B. Martín, B. Aguado and A. Casenave for their excellent technical assistance.

## REFERENCES

- Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T. & Brunak, S. (2005). Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**, 167.
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H. & other authors (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147.
- Berks, B. C. (1996). A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**, 393–404.
- Bourn, W. R. & Babb, B. (1995). Computer assisted identification and classification of streptomycete promoters. *Nucleic Acids Res* **23**, 3696–3703.
- Clayton, T. M. & Bibb, M. J. (1990). *Streptomyces* promoter-probe plasmids that utilise the *xylE* gene of *Pseudomonas putida*. *Nucleic Acids Res* **18**, 1077.
- Demain, A. L. & Fang, A. (2000). The natural functions of secondary metabolites. *Adv Biochem Eng Biotechnol* **69**, 1–39.
- Fernández-Ábalos, J. M., Sánchez, P., Coll, P. M., Villanueva, J. R., Pérez, P. & Santamaría, R. I. (1992). Cloning and nucleotide sequence of *celA1*, and endo- $\beta$ -1,4-glucanase-encoding gene from *Streptomyces halstedii* JM8. *J Bacteriol* **174**, 6368–6376.
- García-González, M. D., Martín, J. F., Vigal, T. & Liras, P. (1991). Characterization, expression in *Streptomyces lividans*, and processing of the amylase of *Streptomyces griseus* IMRU 3570: two different amylases are derived from the same gene by an intracellular processing mechanism. *J Bacteriol* **173**, 2451–2458.
- Geng, D., Baker, D. P., Foley, S. F., Zhou, C., Stieglitz, K. & Roberts, M. F. (1999). A 20-kDa domain is required for phosphatidic acid-induced allosteric activation of phospholipase D from *Streptomyces chromofuscus*. *Biochim Biophys Acta* **1430**, 234–244.
- Ghorbel, S., Kormanec, J., Artus, A. & Virolle, M. J. (2006). Transcriptional studies and regulatory interactions between the *phoR-phoP* operon and the *phoU*, *mtpA*, and *ppk* genes of *Streptomyces lividans* TK24. *J Bacteriol* **188**, 677–686.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.

- Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. & Hirakawa, M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* **34**, D354–D357.
- Kieser, T., Bibb, M., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich, UK: The John Innes Foundation.
- Mansouri, K. & Piepersberg, W. (1991). Genetics of streptomycin production in *Streptomyces griseus*: nucleotide sequence of five genes, *strFGHIK*, including a phosphatase gene. *Mol Gen Genet* **228**, 459–469.
- Martín, J. F. (2004). Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: an unfinished story. *J Bacteriol* **186**, 5197–5201.
- Martín, J. F. & Demain, A. L. (1977). Cleavage of adenosine-5'-monophosphate during uptake by *Streptomyces griseus*. *J Bacteriol* **132**, 590–595.
- Martín, J. F. & Demain, A. L. (1980). Control of antibiotic synthesis. *Microbiol Rev* **44**, 230–251.
- Martín, J. F., Marcos, A. T., Martín, A., Asturias, J. A. & Liras, P. (1994). Phosphate control of antibiotic biosynthesis at the transcriptional level. In *Phosphate in Microorganisms: Cellular and Molecular Biology*, pp. 140–147. Edited by A. Torriani-Gorini, E. Yagil & S. Silver. Washington, DC: American Society for Microbiology.
- Martín, J. F., Gutiérrez, S. & Aparicio, J. F. (2000). Secondary metabolites. In *Encyclopedia of Microbiology*, vol. 4, 2nd edn, pp. 213–236. Edited by J. Lederberg. San Diego, CA: Academic Press.
- Martínez-Domínguez, B., Ibáñez-Gómez, M. V. & Rincón-León, F. (2002). Phytic acid: nutritional aspects and analytical implications. *Arch Latinoam Nutr* **52**, 219–231.
- Mendes, M. V., Tunca, S., Antón, N., Recio, E., Sola-Landa, A., Aparicio, J. F. & Martín, J. F. (2007). The two-component *phoR-phoP* system of *Streptomyces natalensis*: inactivation or deletion of *phoP* reduces the negative phosphate regulation of pimaricin biosynthesis. *Metab Eng* **9**, 217–227.
- Moura, R. S., Martín, J. F., Martín, A. & Liras, P. (2001). Substrate analysis and molecular cloning of the extracellular alkaline phosphatase of *Streptomyces griseus*. *Microbiology* **147**, 1525–1533.
- Patek, M., Muth, G. & Wohlleben, W. (2003). Function of *Corynebacterium glutamicum* promoters in *Escherichia coli*, *Streptomyces lividans*, and *Bacillus subtilis*. *J Biotechnol* **104**, 325–334.
- Rodríguez-García, A., Ludovice, M., Martín, J. F. & Liras, P. (1997). Arginine boxes and the *argR* gene in *Streptomyces clavuligerus*: evidence for a clear regulation of the arginine pathway. *Mol Microbiol* **25**, 219–228.
- Rodríguez-García, A., Barreiro, C., Santos-Beneit, F., Sola-Landa, A. & Martín, J. F. (2007). Genome-wide transcriptomic and proteomic analysis of the primary response to phosphate limitation in *Streptomyces coelicolor* M145 and in a  $\Delta$ *phoP* mutant. *Proteomics* **7**, 2410–2429.
- Schaerlaekens, K., Lammertyn, E., Geukens, N., De Keersmaeker, S., Anne, J. & Van Mellaert, L. (2004). Comparison of the Sec and Tat secretion pathways for heterologous protein production by *Streptomyces lividans*. *J Biotechnol* **112**, 279–288.
- Sola-Landa, A., Moura, R. S. & Martín, J. F. (2003). The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in *Streptomyces lividans*. *Proc Natl Acad Sci U S A* **100**, 6133–6138.
- Sola-Landa, A., Rodríguez-García, A., Franco-Domínguez, E. & Martín, J. F. (2005). Binding of PhoP to promoters of phosphate regulated genes in *Streptomyces coelicolor*: identification of PHO boxes. *Mol Microbiol* **56**, 1373–1385.
- Stieglitz, K., Seaton, B. & Roberts, M. F. (1999). The role of interfacial binding in the activation of *Streptomyces chromofuscus* phospholipase D by phosphatidic acid. *J Biol Chem* **274**, 35367–35374.
- Stieglitz, K. A., Seaton, B. A. & Roberts, M. F. (2001). Binding of proteolytically processed phospholipase D from *Streptomyces chromofuscus* to phosphatidylcholine membranes facilitates vesicle aggregation and fusion. *Biochemistry* **40**, 13954–13963.
- von Döhren, H. & Gräfe, H. (1997). General aspects of secondary metabolism. In *Biotechnology*, vol. 7. *Products of Secondary Metabolism*, pp. 1–55. Edited by H. J. Rehm & G. Reed. Weinheim: VCH Verlag.
- Widdick, D. A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschröder, M. & Palmer, T. (2006). The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **103**, 17927–17932.
- Yang, H. & Roberts, M. F. (2002). Cloning, overexpression, and characterization of a bacterial  $\text{Ca}^{2+}$ -dependent phospholipase D. *Protein Sci* **11**, 2958–2968.

---

Edited by: J. Anné